

Scalable Processor

Abstract

A method and apparatus for issuing and executing memory instructions from a computer system so as to (1) maximize the number of requests issued to a highly pipe-lined memory, the only limitation being data dependencies in the program and (2) avoid reading data from memory before a corresponding write to memory. The memory instructions are organized to read and write into memory, by using explicit move instructions, thereby avoiding any data storage limitations in the processor. The memory requests are organized to carry complete information, so that they can be processed independently when memory returns the requested data. The memory is divided into a number of regions, each of which is associated with a fence counter. The fence counter for a memory region is incremented each time a memory instruction that is targeted to the memory region is issued and decremented each time there is a write to the memory region. After a fence instruction is issued, no further memory instructions are issued if the counter for the memory region specified in the fence instruction is above a threshold. When a sufficient number of the outstanding issued instructions are executed, the counter will be decremented below the threshold and further memory instructions are then issued.